

Tilburg University

Moral hazard and private monitoring

Bhaskar, V.; van Damme, E.E.C.

Published in:
Journal of Economic Theory

Publication date:
2002

Document Version
Peer reviewed version

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Bhaskar, V., & van Damme, E. E. C. (2002). Moral hazard and private monitoring. *Journal of Economic Theory*, 102(1), 16-39.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Moral Hazard and Private Monitoring^{*}

V. Bhaskar
Dept. of Economics
University of Essex
Wivenhoe Park
Colchester CO4 3SQ, UK.
Email: vbhas@essex.ac.uk.

Eric van Damme
CentER for Economic Research
P.O.Box 90153
5000 LE Tilburg
Netherlands
Email: Eric.vanDamme@kub.nl.

^{*}This paper incorporates earlier work by Bhaskar [4] and unpublished notes by van Damme. We are grateful to Tilman Börgers, Dilip Mookherjee, Debraj Ray, an anonymous referee, an associate editor and numerous seminar audiences for useful comments. The first author thanks the CentER for Economic Research (Tilburg) for its hospitality while some of this research was carried out.

Running Head: MORAL HAZARD & PRIVATE MONITORING

Corresponding Author:

V. Bhaskar

Department of Economics

University of Essex

Colchester CO4 3SQ, UK

Email: vbhas@essex.ac.uk

Tel: 44-1206-872744, Fax: 44-1206-872724.

Abstract

We clarify the role of *mixed strategies* and *public randomization (sunspots)* in sustaining near-efficient outcomes in repeated games with private monitoring. In a finitely repeated game where the stage game has multiple Nash equilibria, mixed strategies can support partial cooperation, but cannot approximate full cooperation even if monitoring is “almost perfect”. Efficiency requires extensive form correlation, where strategies can condition upon a sunspot at the end of each period. For any finite number of repetitions, we approximate the best equilibrium payoff under perfect monitoring, assuming that the noise in monitoring is small and sunspots are available. *Journal of Economic Literature* Classification Numbers: C73, D82.

Keywords: repeated games, private monitoring, mixed strategies.

1 Introduction

Repeated games with imperfect public monitoring are well understood. An example is Green and Porter's [13]¹ analysis of a homogeneous good oligopoly, where the individual firm's output is unobserved by its rivals, and the common market price is publicly observed. In a collusive equilibrium, all firms comply with the mandated output reduction. Nevertheless, punishments are triggered after shocks which are sufficiently unfavorable, and hence agents incur payoff losses which may be attributable to the imperfectness of monitoring. These costs are small provided that the signals allow the statistical identification of the deviator and players are patient, as is demonstrated by the "folk-theorem" for this class of games (see Fudenberg, Levine and Maskin [12]).

Rather less is known about repeated games where individuals monitor other players via *private* signals. An example is buyer-seller interaction, where the quality of the product depends stochastically upon the cost or effort incurred by the seller, and where the seller only observes his effort while the buyer only observes the quality he receives. Note the crucial difference with the "standard" model of Klein-Leffler [19] in which the seller's effort is unobservable to the buyer, but where both the buyer and seller know the quality the buyer receives. With private signals, players do not have common knowledge of whether cooperation is to continue or a punishment phase is to be started. Specifically, in the buyer-seller example, the buyer may be reluctant to quit the relationship when he observes bad quality. Since the seller does not observe the quality that the buyer receives, the buyer cannot be sure that the seller will not continue to invest in the relationship. This absence of common knowledge of the players' continuation strategies creates formidable problems, and has deterred the construction of a general theory of such games.

This paper analyzes a simple model of repeated bilateral trade with moral hazard and private monitoring, and highlights the importance of *mixed strategies* and *public randomization* in sustaining near-efficient outcomes. We consider a finitely repeated interaction where the stage game has multiple equilibria. Two traders may supply each other a good of high quality or of low quality. Each trader's action (i.e. the quality supplied) is private information. Moreover, the quality of the good which is received by the recipient is also private. We assume that each trader must incur a sunk cost in order to trade, which gives rise to multiple pure strategy equilibria in the one-shot

¹Abreu, Pearce and Stacchetti [1] provide a general framework for the analysis for this class of games.

trading game. There is one equilibrium where both traders incur the sunk cost and supply low quality, which Pareto dominates the second equilibrium in which neither incurs the sunk cost and no trade takes place. Suppose that this game is repeated twice, and focus on the sustainability of the “efficient outcome” where each trader supplies high quality in period one, and low quality in period two if he receives high quality in period one, but chooses not to trade if he receives low quality. With independent private signals, it is easy to see that this pure strategy profile is not an equilibrium. The essential problem is that in any pure strategy profile, each player chooses a pure action in period one, and hence a player’s beliefs about his opponent’s second period behavior do not vary with the signal he observes. Consequently, it is not optimal for him to punish when he receives a bad signal. This argument extends to the case of correlated signals, provided that the degree of correlation is sufficiently small. Punishments can however be sustained via mixed strategies. In such a mixed strategy equilibrium, a player is uncertain about the pure strategy that his opponent is playing, and he can hence learn about his opponent’s continuation strategy from his signal. This makes it optimal for a player to punish in the event that he observes a bad signal, and hence such a mixed strategy equilibria can support partial cooperation. However, mixed strategies cannot approximate the efficient payoff even if the noise in the signals tends to zero.

This inefficiency arises due to a subtle reason. A player will be willing to punish a bad signal by not trading only if such a signal signifies that his opponent is unlikely to trade tomorrow. If signals are sufficiently uncorrelated, the player will have such beliefs only if the player’s opponent plays, with positive probability, a “bad” pure strategy which sends low quality in period one and does not trade in period two. In other words, cooperation requires defection, not merely in the punishment phase (as in the public signals case), but in the first period itself. Since monitoring is near-perfect, the act of defection is almost surely detected and punished, and hence such a defector’s payoff is bounded away from the efficient payoff. Since defection occurs with positive probability in equilibrium, a player must be indifferent between cooperating and defecting, and hence his overall payoff must equal the payoff from defection. In other words, the set of equilibrium payoffs under private monitoring is bounded away from efficiency even if monitoring is almost perfect.

Efficiency can be restored if players observe the output of a public randomization device (a sunspot) at the end of period 1. Such public randomization, after period one actions have been chosen, allows the players to reduce the severity of punishments, by “forgetting” about past deviations in a coordinated way. By judiciously mitigating the punishment, while still pre-

serving the incentive to cooperate in period one, one can ensure approximate efficiency. We extend this argument to show that in any finitely repeated trading game with imperfect private monitoring, one can approximate the best symmetric equilibrium payoff under perfect monitoring providing that the noise vanishes.

The layout of the remainder of this paper is as follows. Section 2 introduces the basic two period example, and shows that cooperation cannot be sustained by a pure strategy equilibrium. It also that mixed strategies can ensure partial but not full cooperation, while for public randomization ensures approximate efficiency. Section 3 extends these results to the case of any finitely repeated interaction. The final section reviews the related literature and concludes.

2 The Basic Model

Consider the following situation of bilateral trade with moral hazard. Two traders are exchanging goods of variable quality — to make things concrete, think of these as different types of fruit. Each trader must independently make a preliminary investment, incurring a sunk cost F if they are to have the option to trade. If both traders pay this cost, they may proceed to trade. A trader can *cooperate* (action C) by sending fruit of high quality to the recipient, or *defect* (action D), by sending fruit of low quality. High quality fruit has value V_H , which is greater than the value of low quality fruit, V_L . However, the cost of high quality to the supplier, C_H , exceeds the cost of low quality, C_L . Payoffs as a function of the quality dispatched by the trader (which we shall call the *action*, and is indicated by upper case letters) and the quality of fruit received (which we call the *signal*, and indicate using lower case letters) are shown in Fig.1. $A_i = \{C, D, E\}$ is the set of actions of player i and $\Omega_i = \{c, d, e\}$ is the set of possible signals received by i . E is the no-trade option, and we have normalized payoffs by adding the sunk cost to each entry. (It is assumed that if partner i decides not to trade, the other trader j is informed that E has been chosen — we denote this by saying that j receives the signal e for sure, if and only if player i chooses E). In this event, if j has chosen to trade, he loses the sunk cost but not any additional production cost.)

	c	d	e
C	$V_H - C_H$	$V_L - C_H$	0
D	$V_H - C_L$	$V_L - C_L$	0
E	F	F	F

Fig. 1

Assume that if one trader sends the other good fruit, there is a small probability, ϵ , that the fruit deteriorates en route, so that the latter receives low quality, i.e. the recipient gets the signal c with probability $(1 - \epsilon)$ and the signal d with probability ϵ . If the sender sends bad quality, the receiver gets bad quality (signal d) for sure, and if the a trader chooses action E , the other trader gets signal e for sure. We may then write the stage game payoffs as in Fig.2, which shows the payoff to the row player. $\tilde{V}_H = (1 - \epsilon)V_H + \epsilon V_L$ is the “expected quality” received when high quality is dispatched.

	C	D	E
C	$\tilde{V}_H - C_H$	$V_L - C_H$	0
D	$\tilde{V}_H - C_L$	$V_L - C_L$	0
E	F	F	F

Fig. 2: The Game G

Assume that it is efficient to both traders to exchange high quality fruit, so that $\tilde{V}_H - C_H > V_L - C_L$. Quality dispatched and quality received are both unverifiable, and hence high quality trade cannot be legally enforced. Clearly, the action C is strictly dominated. However, both (D, D) and (E, E) are Nash equilibria of the game G , and there is also a mixed Nash equilibrium where each trader plays D with probability $\mu^* = \frac{F}{V_L - C_L}$ and E with probability $1 - \mu^*$. Assume that low quality trade is sufficiently better than no-trade so that $V_L - C_L - F > \Delta C = C_H - C_L$, and focus attention on the case where G is played twice.²

Suppose that each player cannot observe the quality dispatched by the other player, i.e. *actions* are unobserved. The central focus of this paper is on the case where each player’s signal is private, i.e. he only knows what quality he received. However, to provide a benchmark, we first briefly discuss the case analyzed in the literature, when the quality received by any trader is commonly observed, i.e. the signals are *public*. Since G has multiple Nash equilibria, we may, as in Benoit and Krishna [3], construct an equilibrium where C is played in period one. Each player adopts the following strategy: choose C in period one; in period two, play D if the signals are (cc) , and play E otherwise. To see that this strategy profile is an equilibrium, note that in period two, each player knows the action that his opponent will play for sure, and hence his own action is optimal at every information set. Given second

²If trade is seasonal, as is likely in the fruit example, the finitely repeated game may be a better representation of interaction than an infinitely repeated game. In addition, the two-period example allows us to characterize the efficiency properties of *all* equilibria.

period behavior, a deviation to D in period one is unprofitable. Equilibrium payoffs are given by

$$\tilde{V}_H - C_H + (1 - \epsilon)^2(V_L - C_L) + (1 - (1 - \epsilon^2))F$$

This payoff is lower than the *efficient payoff* of $(\tilde{V}_H - C_H + V_L - C_L)$, which is an equilibrium payoff if the players actions were to be observed.³ Imperfect monitoring via public signals creates an inefficiency relative to the efficient payoff, but this inefficiency is of order ϵ , and vanishes as ϵ tends to zero.

Consider now an alternative information structure which is the focus of this paper, where each trader observes the quality of fruit he receives but does not observe the quality received by the other trader — signals are *private*. Hence neither the quality sent nor the quality received by trader i are mutual knowledge between the traders, although they could be arbitrarily close to being so if the noise (ϵ) is small. It is convenient to be slightly more general with respect to the signalling technology and to allow for correlation between the players' signals conditional on the action taken. For $a \in A = A_1 \times A_2$, let $\omega = (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$ be the profile of signals realized, where player i observes only ω_i . Furthermore, assume that conditional on $a = (C, C)$, the signal distribution is given by

		Trader 2's signal	
		c	d
Trader 1's signal	c	$(1 - \epsilon)^2 + \rho\epsilon(1 - \epsilon)$	$(1 - \rho)\epsilon(1 - \epsilon)$
	d	$(1 - \rho)\epsilon(1 - \epsilon)$	$\epsilon^2 + \rho\epsilon(1 - \epsilon)$

Fig. 3 Distribution of signals conditional on (C, C)

If $a = (C, D)$, $\omega = (c, d)$ with probability $1 - \epsilon$, and $\omega = (d, d)$ with probability ϵ . If $a = (C, D)$, $\omega = (c, d)$ with probability $1 - \epsilon$, and $\omega = (d, d)$ with probability ϵ . If $a = (D, D)$, $\omega = (d, d)$ with probability one. If $a = (E, E)$, $\omega = (e, e)$ with probability one, and if player i chooses E and if player j chooses either C or D , then $\omega_j = e$ and player i is informed that $\omega_i \in \{c, d\}$.

This signalling structure is parametrized by ϵ and ρ , where ϵ is the level of “noise”, and ρ is the degree of conditional correlation between signals (conditional on the action profile (C, C)).⁴ Since all probabilities must be

³We call this the efficient payoff since this is the maximum payoff that each player can achieve in any equilibrium.

⁴For simplicity we shall call ρ the degree of correlation, by which we mean *conditional correlation*, i.e. conditional upon the action profile (C, C) . In our model, correlation could arise due to correlated weather shocks which affect the quality received by both players.

positive, we must have that $\rho \leq 1$ and $\rho \geq \max\{-\frac{\epsilon}{1-\epsilon}, -\frac{1-\epsilon}{\epsilon}\}$. Assume that ρ satisfies these inequalities strictly, thus ensuring that all signal combinations have positive probability when (C, C) is played. $\rho = 0$ corresponds to case where the signals are independent, while if $\rho = 1$, the signals are perfectly positively correlated — this is equivalent to the public signals case.

2.1 Pure Strategy Equilibria

Our focus is on the twice repeated game, which we denote $G^2(\epsilon, \rho)$. Players maximize the sum of expected payoffs in the two stages. A pure strategy for a player i in $G^2(\epsilon, \rho)$ is a pair $s_i = (f_i, g_i)$ where $f_i \in A_i$ is the action taken in the first period, and $g_i : A_i \times \Omega_i \rightarrow A_i$ specifies the action taken in the second period as a function of the player's first period action and the signal he receives. Our focus is on the sequential equilibria of $G^2(\epsilon, \rho)$.⁵

Consider first the case where signals are independent, so that $\rho = 0$. In this case we cannot support the playing of C in period one in any pure strategy equilibrium, even if ϵ is arbitrarily small. Suppose that C is chosen by both traders in period one. This can only be optimal for each trader if he believes that the other trader will reward signal c and punish signal d . Hence each player's strategy must be of the type: play C in period 1; in period two, play D on receiving signal c , and play E on receiving signal d .⁶ However, such a strategy is not a best response to itself; it is not optimal for a trader who receives signal d to carry out this punishment. Suppose that I am a player who believes that my opponent is playing such a strategy. If I observe the signal d , I should attribute this to the error in the signalling technology — the application of Bayes' rule to my opponent's strategy implies that this is the only event which has positive probability. Since I have chosen C in period one, I know that my opponent will receive signal c with very high probability, $1 - \epsilon$. Hence it is optimal for me to continue with D , and ignore the signal I have received. Since varying second period behavior with the first period signal is not optimal, this makes it impossible to support the

⁵Our focus is on efficient equilibria, i.e. on strategy profiles where C is played in period one, and in this case signals c and d will both be observed with positive probability. In consequence, we could as well use the Nash equilibrium criterion, since this will require optimal behavior at all information sets which are reached.

⁶A player could also punish by playing the mixed equilibrium, but the argument which follows also applies in this case.

playing of C with probability one in the first period.^{7 8}

Consider now the case of correlated signals. If this correlation is positive, if a player receives a bad signal, this makes it more likely that his opponent has also received a bad signal. Consequently an agreement to punish on receiving a bad signal could be made self enforcing. However, the degree of positive correlation must be large enough. Define the strategy α as follows

Strategy α : 1st period: C . 2nd period: D if (Cc) , E otherwise.

Consider the sustainability of the strategy profile (α, α) .⁹ To check that this is a Nash equilibrium we need to see that second period behavior is optimal. If my opponent is playing the strategy α , then he will play D in period 2 if he has observed the signal c , and will play E if he has observed d . Hence conditional on my first period action C , and on my receiving the signal c , the probability that my opponent plays D in period 2, $\mu_i(Cc; \alpha)$ equals $(1 - \epsilon) + \rho\epsilon$. Similarly, conditional on my first period action C , and on my receiving the signal d , the probability that my opponent plays D in period 2, $\mu_i(Cd; \alpha)$ equals $(1 - \rho)(1 - \epsilon)$. Since α requires me to play D on observing c and E on observing d , I must believe that my opponent plays D in the former event with probability greater than μ^* , and with probability less than μ^* in the latter event (recall that μ^* is the probability with which D is played in the mixed equilibrium of G). I.e. we must have

$$\mu_i(Cc; \alpha) = (1 - \epsilon) + \rho\epsilon \geq \mu^*. \quad (1)$$

$$\mu_i(Cd; \alpha) = (1 - \rho)(1 - \epsilon) \leq \mu^* \quad (2)$$

In addition, it must be optimal to play C in period one, rather than deviating by playing D in period one and E in period two. Let $\Delta C =$

⁷This argument appears to be quite general — given independent signals where every signal has positive probability under any action profile, and generic payoffs in the stage game, the pure strategy equilibria of the twice repeated game must be degenerate, i.e. repetitions of stage-game Nash equilibria. By using induction, this result may also be extended to any number of finite repetitions.

⁸One possible solution to this coordination problem is to allow players to communicate at the end of period one. This route is explored by Compte [9] and Kandori and Matsushima [18], who use this to prove versions of the folk theorem of for infinitely repeated games with private monitoring. The focus of the present paper is a purely non-cooperative analysis, without communication. Nevertheless, it may be worth pointing out that in the present finitely repeated game, communication is ineffectual unless signals are sufficiently highly correlated — see Bhaskar ([7]) for details.

⁹Any pure strategy equilibrium where C is played in period one must be similar to (α, α) , since signal c must be rewarded and d must be punished. What happens after signal e is irrelevant.

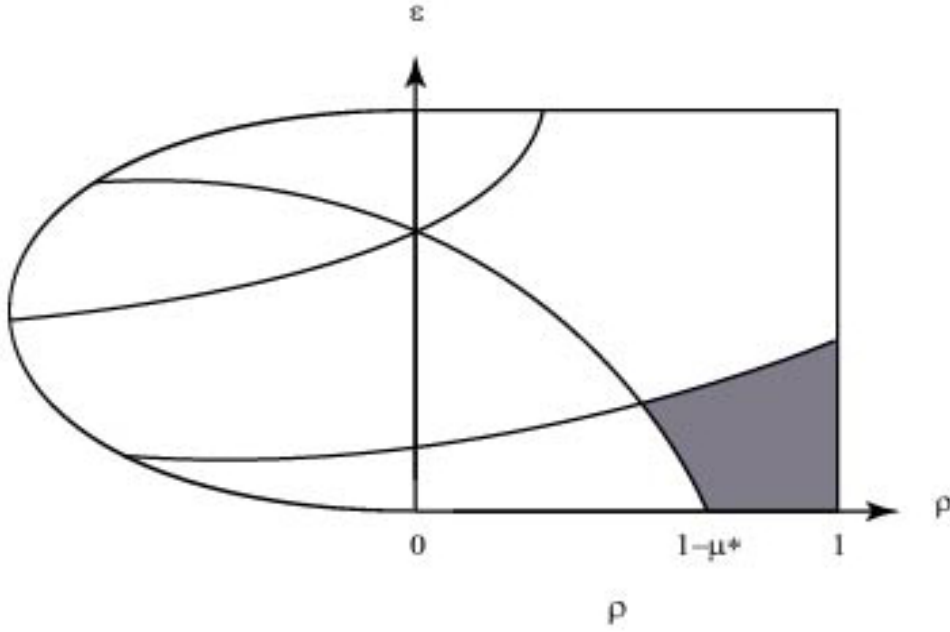


Fig. 4: Region of Sustainable Co-operation

$C_H - C_L$ — this is the first period gain to deviating by producing low quality. This must be less than the second period loss from deviation, i.e.

$$\Delta C \leq [((1 - \epsilon)^2 + \rho\epsilon(1 - \epsilon))(V_L - C_L) + \epsilon F] - F \quad (3)$$

Inequalities (1-3) are graphed in Fig. 4. The shaded area in this figure shows values of ϵ and ρ such that these inequalities are satisfied, and (α, α) is an equilibrium. The key features of this figure are summarized in the following proposition.

Proposition 1 *i) If $\rho \geq 1 - \mu^*$, cooperation can be supported by a pure strategy equilibrium if ϵ is sufficiently small.*

ii) If $\rho < 1 - \mu^$, cooperation cannot be supported by a pure strategy equilibrium if ϵ is sufficiently small.*

iii) If ρ is close to but less than $1 - \mu^$, cooperation can be supported if ϵ is neither too large nor too small.*

Note that correlation must be sufficiently high for cooperation to be supported. Most intriguing is part (iii) of the proposition, on the relation between the level of noise and cooperation at intermediate levels of correlation. (2) will not hold if ϵ is small and close to zero, but Fig. 4 shows that this inequality can be satisfied for larger values of ϵ . However, ϵ must not be too large since otherwise (3) will not be satisfied. Hence the set of pure strategy equilibrium outcomes is not monotone in ϵ .¹⁰

We shall henceforth focus attention upon the case where $\rho < 1 - \mu^*$, when cooperation cannot be sustained via pure strategies. We refer the interested reader to Mailath and Morris [20], who discuss correlated signals in greater detail and prove a folk theorem for infinitely repeated games with private signals if these signals are sufficiently highly correlated.

2.2 Mixed Strategies

We now construct a mixed strategy equilibrium which allows us to support partial cooperation in the twice repeated game for any level of correlation between signals. A mixed strategy for a player i is a probability vector σ_i , where $\sigma_i(s_i)$ denotes the probability assigned to the pure strategy s_i . Note that we shall conduct our analysis in terms of mixed strategies rather than behavior strategies.

In order to understand the role of mixed strategies, it is useful to interpret the reason why pure strategies are unable to support any cooperation. For intuition, focus on the case where signals are independent so that $\rho = 0$. Observe that in this case, from (1) and (2) that $\mu_i(Cc, \alpha) = \mu_i(Cd, \alpha) = 1 - \epsilon$. In other words, if a player “knows” his opponent’s strategy (as is implicit in a pure strategy equilibrium), his beliefs regarding his opponent’s action in period two depend only upon his prior knowledge, and are insensitive to the signal he receives. To make a player willing to respond to the signal, we must ensure that it conveys some information about his opponent’s second period

¹⁰Proposition 3 below shows that payoffs in any equilibrium are bounded away from the efficient payoff if $\rho < 1 - \mu^*$. Hence the paradoxical finding, that equilibrium payoffs are not monotone in ϵ , applies even if we consider mixed strategies — see the remarks at the end of section 3.

actions in equilibrium. More specifically, a player will be willing to respond differently to different signals only if these signals indicate that his opponent is likely to play differently.¹¹ This is possible if we allow for mixed strategies, since the player's prior beliefs will not be degenerate, and the signal allows him to learn which pure strategy his opponent is playing.

Consider the following pure strategies for the repeated game.

Strategy α : 1st period: C . 2nd period: D if (Cc) , E otherwise.

Strategy β : 1st period: D . 2nd period: E .

The payoff matrix for these two supergame strategies is:

	α	β
α	$\tilde{V}_H - C_H + V_L - C_L - \Gamma(\epsilon)$	$V_L - C_H + F$
β	$\tilde{V}_H - C_L + F$	$V_L - C_L + F$

where $\Gamma(\epsilon) = [1 - (1 - \epsilon)^2 - \rho\epsilon(1 - \epsilon)](V_L - C_L) + \epsilon F$ is a term of order ϵ .

Confining attention to the pure strategy set $\{\alpha, \beta\}$ for each player, we see that α is a strict best response to α if ϵ is sufficiently small and β is a strict best response to β . Hence the above payoff matrix also has a symmetric mixed strategy equilibrium where each player plays α with probability π and β with probability $1 - \pi$, where $\pi = \frac{\Delta C}{V_L - C_L - F - \Gamma(\epsilon)}$. Call this mixed strategy $\hat{\sigma}$. We now show that the symmetric strategy profile $(\hat{\sigma}, \hat{\sigma})$ is an equilibrium of the repeated game.

Proposition 2 *The symmetric strategy profile where each player plays $\hat{\sigma}$ is an equilibrium of $G^2(\epsilon, \rho)$ for any $\rho < 1$ if ϵ is sufficiently small.*

Proof. Assume that the opponent plays $\hat{\sigma}$. It is easily seen that any strategy that starts by playing E is strictly inferior. Write $\mu_i(\cdot; \hat{\sigma})$ for the beliefs induced by $\hat{\sigma}$, i.e. the probability that the opponent will play D at $t = 2$. Then

$$\mu_i(Cc; \hat{\sigma}) \rightarrow 1 \text{ as } \epsilon \rightarrow 0 \quad (4)$$

$$\mu_i(Cd; \hat{\sigma}) \rightarrow 0 \text{ as } \epsilon \rightarrow 0 \quad (5)$$

$$\mu_i(D\omega_i; \hat{\sigma}) = 0 \quad (6)$$

¹¹Alternatively, a player can be made willing respond to the signal even with constant beliefs if $\mu^* = 1 - \epsilon$, so that he is indifferent between his two actions and takes different actions at different information sets. We discuss this possibility, due to Kandori [16] after proposition 2.

At information set (Cc) , I know that my opponent has played α and that he received signal c with probability almost 1, and hence (4) follows. At information set (Cd) , the signal d could have arisen either because (i) my opponent is playing α and the noise intervened, or (ii) my opponent is playing the strategy β . The probability that my opponent continues with D equals the conditional probability that my opponent is at the information set (Cc) given that I am at (Cd) , and equals

$$\mu_i(Cd; \hat{\sigma}) = \frac{\pi\epsilon(1-\epsilon)(1-\rho)}{(1-\pi) + \pi\epsilon} \quad (7)$$

The condition that $\mu_i(Cd; \hat{\sigma}) \leq \mu^*$ is equivalent to the condition that

$$\pi \leq \pi^* = \frac{\mu^*}{(1-\epsilon)[\mu^* + \epsilon(1-\rho)]} \quad (8)$$

Since $\pi^* \rightarrow 1$ as $\epsilon \rightarrow 0$ while $\pi \rightarrow \frac{\Delta C}{V_L - C_L - F} < 1$ as $\epsilon \rightarrow 0$, we will have $\mu_i(Cd; \hat{\sigma}) \leq \mu^*$ as long as ϵ is sufficiently small — indeed, (5) also follows from this. Finally, (6) follows since the opponent is sure to receive signal d after D and since both α and β play E after d . (4-6) together with the fact that both (D, D) and (E, E) are strict equilibria of G imply that for ϵ small enough, D is the unique best response at (Cc) , and E is the unique best response at other information sets at $t = 2$. It follows that both α and β prescribe best responses to $\hat{\sigma}$ at $t = 2$. Since, by construction, α and β are also best responses at $t = 1$, $(\hat{\sigma}, \hat{\sigma})$ is an equilibrium of the game. ■

The above construction is very different from Kandori's early work[16]. Kandori analyzes a twice repeated game where the stage game that has a unique mixed strategy equilibrium, and the private signals are independent. Kandori constructs an equilibrium where the efficient action profile is played with probability one in period one. Since the signals are independent, a player will have the same beliefs about his opponent's actions in period two after any private signal, and these beliefs are constructed to correspond to the mixed equilibrium of the stage game. However, since a player is indifferent between all pure actions in the support of the mixed strategy equilibrium, he will be willing to play different continuation strategies in response to these signals, thus providing incentives for cooperative behavior in period one.

In Kandori's equilibrium, a player has identical beliefs about his opponent's continuation strategy at different histories, but chooses his continuation strategy in period two differently depending upon which history materializes. We argue that this equilibrium is not robust, for the following reason. Suppose that each player's stage game payoffs are subject to a small amount of incomplete information, as in Harsanyi [15]. In this case, at stage 2, a

player will behave in the same way after different histories for almost any realization of his payoff information, since he will strictly prefer one action above the other. In other words, this equilibrium cannot be purified in the manner of Harsanyi, if we perturb stage game payoffs, or equivalently, assume that payoffs in the perturbed repeated game are additively separable.¹² This criticism does not apply to the equilibrium we have constructed: a player is required to randomize only at stage 1 and has strict incentives to follow the recommendations of his strategy at stage 2. Since the equilibrium strategy is measurable with respect to a player's beliefs, it is not difficult to construct equilibria of incomplete information games that approximate it.

Although the mixed strategy equilibrium supports partial cooperation, the probability with which the players play C in period one is bounded away from one even if ϵ is arbitrarily small. To see this, observe that π , the probability with which the strategy α is played, tends to $\frac{\Delta C}{V_L - C_L - F} < 1$ as $\epsilon \rightarrow 0$. Hence the equilibrium payoff in the game without any noise cannot be approximated by this mixed equilibrium, in contrast with the situation where signals are publicly observed. We now show that this result holds more generally — the cooperative equilibrium under perfect monitoring cannot be approximated under imperfect monitoring even if the noise in the signals goes to zero. In other words, the sequential equilibrium outcome correspondence is not lower-hemicontinuous.¹³

Proposition 3 *If ρ is fixed and strictly less than $1 - \mu^*$, the efficient outcome where both traders produce high quality in period one, and low quality in period two cannot be approximated by any equilibrium of $G^2(\epsilon, \rho)$, as $\epsilon \rightarrow 0$.*

Proof. Assume that $(\sigma_1(\epsilon), \sigma_2(\epsilon))$ is a mixed Nash equilibrium of $G^2(\epsilon, \rho)$ that is approximately efficient, i.e. each player's payoff is approximately $\tilde{V}_H - C_H + V_L - C_L$.

Define first the set Θ of *good* pure strategies in the repeated game, where a good strategy plays C in period 1, and responds to the signal c by playing D in period 2. $\Theta = \{(f_i, g_i) : f_i = C \text{ and } g_i(c) = D\}$. If the outcome of any mixed strategy is to be approximately efficient, then both players must be

¹²See Bhaskar [5],[6] for an analysis of such payoff perturbations in the context of repeated games and other dynamic games with additively separable payoffs and private monitoring.

¹³This failure of lower-hemicontinuity is with respect to the information structure, and hence quite different from the example of Radner, Maskin and Myerson [23], who consider the behavior of equilibrium payoffs in a repeated game with public signals as the discount rate tends to one, given a fixed information structure. Kandori [17] discusses the effects of improved information in the case of imperfect public monitoring.

playing good strategies with probability close to one. In this event, neither player will play E in period one, since this yields a strictly lower payoff.

Since player i is playing C or D in period one, player j 's first period payoff gain from playing D rather than C in period 1 equals ΔC . To ensure that player j has an incentive to play C in period one, we must ensure that player j suffers a second period loss of at least ΔC if he plays D in period one. Hence player i must be playing a good strategy which rewards the signal c (by playing D in period two) and punishes the signal d (by playing E in period two). Call any such strategy α — i must assign positive probability to a pure strategy α . Since this argument applies for $i = 1, 2$, α is in the support of both players' strategies.

Let α' be a pure (good) strategy which plays C in period one, and responds to signal d by playing D in period two, i.e. this strategy does not punish after d .

Define the set Ξ of *bad* strategies as follows — any strategy from Ξ plays D in period one, and responds to the signal c by playing E . We now show that if player i assigns positive probability to α , then player j must assign positive probability to a bad strategy. We do this by showing that if no bad strategy is in the support of player j 's mixed strategy, then α is strictly inferior to α' .

Assume that no bad strategy is in the support of player j 's mixed strategy. Note that against $\sigma_j(\epsilon)$, α and α' yield the same expected payoff in the first period, and also in the second period when i receives signal c . Hence, condition on j playing $\sigma_j(\epsilon)$, i playing α or α' and i receiving signal d . There are now two possibilities: player j is playing a pure strategy in the support of $\sigma_j(\epsilon)$ with $f_j = D$ or with $f_j = C$.

In the first case ($f_j = D$), since j is not playing a bad strategy and since he gets c with probability $(1 - \epsilon)$, he is most likely to play D . Consequently, in this case α' yields strictly more than α .

In the case when $f_j = C$, both players chose C in the first period and fig. 3 shows that j received signal c with probability $(1 - \rho)(1 - \epsilon)$. Since j is playing a good strategy with probability close to one, he continues with D with probability close to one after receiving signal c . We therefore conclude that, conditional on i receiving signal d and $f_j = C$, i believes that j will play D with probability approximately $1 - \rho$ or more. Hence if $1 - \rho > \mu^*$, then α' is strictly better than α in this case as well.

We conclude that if j does not play a bad strategy and if $\rho < 1 - \mu^*$, then α' yields strictly more than α when ϵ is sufficiently small. Since α is the support of $\sigma_i(\epsilon)$ for each player i , each player j must be playing a bad strategy with positive probability when $\rho < 1 - \mu^*$.

Now, if j plays a good strategy and i plays a bad one, then i 's payoff is $\tilde{V}_H - C_L + F$, and hence against $\sigma_j(\epsilon)$ a bad strategy yields approxi-

mately $\tilde{V}_H - C_L + F$ in equilibrium. Since the payoff to all pure strategies in $\sigma_i(\epsilon)$ must be equal in any mixed Nash equilibrium, this implies that neither player's payoff can be greater than $\tilde{V}_H - C_L + F$. Since the efficient outcome has a strictly greater payoff, it cannot be approximated by any mixed Nash equilibrium of $G^2(\epsilon, \rho)$, no matter how small ϵ is. ■

Note that this proposition also implies that if ρ is less than but close to $1 - \mu^*$, the mixed strategy equilibrium payoffs are not monotone in ϵ . Cooperation can be sustained via pure strategies for intermediate values of ϵ , but not for ϵ close to zero, since in this case mixed strategies are required.

The basic argument underlying the proof is as follows. If an equilibrium is to be approximately efficient, both players must play *good* pure strategies with high probability, where a good strategy is defined as one which plays C in the first period, and responds to the signal c with D — this is the only way in which the outcome can approximate (C, C) in period one, and (D, D) in period two. Since a strategy which plays D in the first period will have a higher first period payoff against such a good strategy, equilibrium requires that the signal d must be punished. Hence both players must play, with positive probability, the strategy α which plays C in period one, and in period two, punishes signal d by playing E , and rewards signal c by playing D . However, if α is to be optimal for a player, say player i , his beliefs about player j 's continuation strategy must vary sufficiently with the signal he observes. Specifically, the signal d must indicate that player j is likely to play E , even though the signal c indicates that j is likely to play D . If the extent of correlation is small, such variation in i 's beliefs is only possible if j plays with positive probability a strategy which plays D in period one, and responds to signal c with E — we call any such strategy a *bad* strategy. If ϵ is small, j need play a bad strategy only with a small probability in order make i 's beliefs sufficiently responsive. However, if j plays a bad strategy, the payoff of j must be low — for example if i plays a good strategy, then j earns at most a payoff of $\tilde{V}_H - C_L + F$, which is strictly less than the efficient payoff. Now, if a bad strategy is in the support of the player's equilibrium mixed strategy, the player's overall payoff must be *exactly equal* to the payoff produced by the bad strategy. Consequently, equilibrium payoffs are bounded away from efficiency. In contrast, if monitoring is public (or private signals are sufficiently correlated), only good strategies need be played, and inefficiencies are only triggered after unfavorable signals. Hence efficiency is ensured if the noise is sufficiently small.

Two facets of the above argument need emphasis. First, if ϵ is small, j need play a bad strategy only with a small probability, of order ϵ — this is sufficient to make i 's beliefs responsive to his signal. Hence j playing a bad strategy need not (by itself) have a large negative effect on i 's payoffs.

However, when j plays a bad strategy (or more generally, when he plays D in period one), he has to be punished, and this has a large negative effect on j 's payoff. Since a bad strategy is in the support of j 's equilibrium strategy, j 's equilibrium payoff must be inefficient. This argument suggests that if the punishment of a bad strategy can be mitigated judiciously so that the first period gain is just offset by the future loss, one can ensure efficiency. We now show that public randomization provides such a mechanism for mitigating punishments.

2.3 Sunspots & Efficiency

The previous analysis suggests that the key to ensuring efficiency is to soften the punishment meted out to first period defection. How do we soften the punishment to defection? One possibility is that in period two, each player does not always punish the signal d , but merely punishes with some probability, by randomizing between E and D in the event of receiving signal d . However, such randomization at the individual level is infeasible, since each player has strict incentives to play D at this information set. What is required is that player can agree to forget past transgressions in a *coordinated* way. A sunspot, i.e. the realization of a commonly observed random variable, can play this role. Intuitively, players can agree to forget about past transgressions with some probability, so that defectors are deterred, but not too harshly. Formally, the sunspot allows for *extensive-form correlation*, which transforms the base game by convexifying the set of equilibrium payoffs, allowing the two players to achieve any payoff in the interval $[F, V_L - C_L]$. Consequently, a player who chooses D in period one can be punished so that her payoff loss in period 2 is arbitrarily close to her payoff gain in period 1. Since there is no overall payoff loss from playing a bad strategy, this enables both players to play a bad strategy with small probability.

Assume that at the end of period one players can publicly observe the outcome ϕ_1 of a random variable Φ_1 , which is uniformly distributed on $[0, 1]$. The sunspot convexifies the set of equilibrium payoffs of G . Specifically, for any $m \in [0, 1]$, the correlated strategy $z = (z_1, z_2)$ with

$$z_1(\phi_1) = z_2(\phi_1) = \begin{cases} E & \text{if } \phi_1 \leq m \\ D & \text{if } \phi_1 > m \end{cases} \quad (9)$$

is a correlated equilibrium of G . By varying m , any payoff Z in $[F, V_L - C_L]$ can be obtained in this way. Note that such a correlated equilibrium z is strict: if a player believes that his opponent plays z_j with probability greater than $\min\{\mu^*, 1 - \mu^*\}$, then it is optimal to play z_i himself. Let z be such a

correlated equilibrium of G with payoff Z , and modify the strategies α and β from the previous section such that E is replaced by z . The only thing that changes in the payoff matrix is that F has to be replaced by Z . Provided that

$$Z + \Delta C < V_L - C_L - \Gamma(\epsilon) \quad (10)$$

(an inequality which is satisfied for Z sufficiently close to F), (α, α) and (β, β) are still strict equilibria of this 2×2 game, and as in the previous section, there exists a mixed strategy equilibrium of this payoff matrix, where α is played with probability π and β with probability $1 - \pi$. The claim of the previous section, that this is an equilibrium of the repeated game, continues to apply. Observe from the proof that the only essential change occurs when considering the information set (Cd) . For any given π , I attach a probability greater than $\min\{\mu^*, 1 - \mu^*\}$ to my opponent continuing with z_j provided that ϵ is sufficiently small.¹⁴ Since z is strict, it is optimal for me to continue with z_i as well. Now, investigate the consequences of varying Z . By increasing Z towards the upper bound from (10), the probability π can be increased to π^* (cf.(8)). However, $\pi^* \rightarrow 1$ as $\epsilon \rightarrow 0$, and hence the players will play (α, α) with probability close to one, and will obtain a payoff close to the efficient one.

Observe that time at which the output of the public randomization device is observed by both players is crucial. This must be *after* players have chosen their actions in period 1, but before they choose actions in period 2. In other words, *extensive form correlation* is essential. Extensive-form correlation was introduced by Myerson [22], who also pointed out this allows greater strategic possibilities than normal-form correlation.

3 Many Repetitions

We now consider an arbitrary finite number (T) of repetitions of the stage game with imperfect monitoring. Our object is to show that if ϵ is sufficiently small, one can approximate the maximal symmetric equilibrium payoff under perfect monitoring, $V^*(T)$, which is defined by:

$$V^*(T) = \frac{T-1}{T}(\tilde{V}_H - C_H) + \frac{1}{T}(V_L - C_L) \quad (11)$$

First we show that in order to obtain a general efficiency result, one must allow for players to condition their actions upon a public randomization device. The result relies on an adaptation of the argument of proposition 3,

¹⁴The relevant condition is inequality (7), which specifies how small ϵ must be given π .

applied to the last two periods of the T period game. However, it is not immediate since the private signals in previous periods allow some endogenous correlation of strategies — the continuation strategies in the final two periods correspond therefore to a correlated equilibrium of the two period game. Under the hypotheses of the proposition, this correlation is insufficient, so that inefficient randomization is required in the penultimate period.

Proposition 4 *If players cannot condition their actions upon a sunspot, and $\mu^* < \frac{1-\rho}{2-\rho}$, the efficient payoff $V^*(T)$ cannot be approximated by any equilibrium of the T period repeated game, as $\epsilon \rightarrow 0$.*

Proof. Approximate efficiency requires that the path where (C, C) is played in the first $T - 1$ periods and (D, D) is played at T is realized with probability close to one. Let \hat{h} be the $T - 2$ period private history where (Cc) is realized in every period. Approximate efficiency requires that at \hat{h} the player must play, with probability close to one, continuation strategies from the set Θ of *good* pure continuation strategies, which play C at $T - 1$, and responds to the signal c with D in period T .

If player j is at \hat{h} , he assigns probability $1 - \epsilon(1 - \rho)$ to player i also being at \hat{h} , and likewise playing a good continuation strategy with probability close to one. Hence player j 's benefit from playing D in period $T - 1$ is approximately $\Delta C > 0$. To ensure that player j has an incentive to play a good continuation strategy, player i must be playing, with positive probability, a good continuation strategy α which rewards the signal c (by playing D) and punishes the signal d (by playing E), in the final period. Since this argument applies for $i = 1, 2$, α is in the support of both players' continuations strategies. Let α' be a good continuation strategy which plays C in period $T - 1$, and responds to signal d by playing D in period T . Define the set Ξ of *bad* continuation strategies as follows – any strategy from Ξ plays D in period $T - 1$, and responds to the signal c by playing E . We now show that under the hypotheses of the proposition, if no bad strategy is in the support of player j 's mixed continuation strategy, then α is strictly inferior to α' .

Condition now on the $T - 1$ period history where \hat{h} is followed by Cd for player i . Since j is not playing a bad continuation strategy at the history \hat{h} , the probability that he plays D in the final period is at least

$$\frac{[1 - \epsilon(1 - \rho)](1 - \rho)(1 - \epsilon)\epsilon}{[1 - \epsilon(1 - \rho)]\epsilon + (1 - \rho)\epsilon} \quad (12)$$

which converges to $\frac{1-\rho}{2-\rho}$ as $\epsilon \rightarrow 0$. Since $\frac{1-\rho}{2-\rho} > \mu^*$, it is optimal to play D at this information set for ϵ sufficiently small and hence α' is strictly better than α . Hence each player must be playing a bad continuation strategy at \hat{h}

with positive probability. If j plays a good continuation strategy and i plays a bad one, then i 's continuation payoff is $(\tilde{V}_H - C_L + F)/T$, and hence his overall payoff cannot approximate $V^*(T)$. ■

We now assume that players can observe a sunspot at the end of each period, and construct an efficient equilibrium. Our construction of the strategy for the T period game, σ^T , is a recursive one, and utilizes the efficient strategy profiles σ^τ for all $\tau < T$. Suppose that a player is playing some strategy σ^τ , in period $t - 1$, where $T \geq \tau \geq t - 1$. His continuation strategy in period t depends upon the realization of the sunspot ϕ_{t-1} at the end of period $t - 1$. If ϕ_{t-1} is less than some critical value, the player continues with the strategy σ^τ . On the other hand, if ϕ_{t-1} is greater than this critical value, the players “forget” all past private information and begin afresh with the efficient repeated game strategy for the r period repeated game, σ^r , where $r = T - (t - 1)$. In other words, the *length* of private history that players condition their behavior on depends upon the sequence of sunspot realizations $(\phi_1, \phi_2, \dots, \phi_{t-1})$. If q is the index of the last time period such ϕ_q was greater than the critical value, then the players will be playing the strategy σ^{T-q} in period t , and conditioning their behavior on private information relating to the last $(t - 1) - q$ periods.

To define any $\tau = T - q$ period strategy, partition the set of $(t - 1) - q$ period private histories into two subsets. Call such a history a *good* history if the player has always played C , if his signal in every period is either c or d , and if the signal at date $t - 1$ was c . The strategy will play D after the signal d so that the only good history which arises on the path of play is (Cc, \dots, Cc) , i.e. one where in each of the last $(t - 1) - q$ periods, the action C has been taken and signal c has been observed.¹⁵ Call any other history a *bad* history — at a bad history, either a player has played D or E or observed signal e in some period, or has observed d at date $t - 1$.

The strategy σ^T is defined as follows:

1. At period 1, play C with probability π' , D with probability $1 - \pi'$.
2. Let $t \in \{2, 3, \dots, T - 2\}$, and suppose that at date $t - 1$ the player was playing the strategy σ^τ , where $T \leq \tau \leq T - (t - 1)$:
 - (a) If $\phi_{t-1} > m'$, play $\sigma^{T-(t-1)}$, the equilibrium strategy in the $T - (t - 1)$ period repeated game. This plays C with probability π' , D with probability $1 - \pi'$ in the current period.

¹⁵A good history of the type (Cc, \dots, Cd, Cc) can also arise when a player deviates from his strategy. Since a player's beliefs about his opponent's continuation strategy are the same at any history which is good, we define σ^T so that the player plays the same continuation strategy at both these good histories.

- (b) If $\phi_{t-1} \leq m'$, play C if the $(t-1) - (T-\tau)$ period history is a good history and play D at any bad history.
3. At period $T-1$ suppose that at $T-2$ player was playing the strategy σ^τ , where $T \leq \tau \leq 3$:
- (a) If $\phi_{T-2} > m_{T-2}$: play σ^2 , the equilibrium strategy in the 2 period repeated game.
- (b) If $\phi_{T-2} \leq m_{T-2}$: if the $(\tau-2)$ period private history is a good history, play C with probability π_{T-1} and D with probability $1 - \pi_{T-1}$; play D at any bad history.
4. At period T , suppose that at date $T-1$ the player was playing the strategy σ^τ , where $T \leq \tau \leq 2$:
- (a) If $\phi_{T-1} > m_{T-1}$: play D .
- (b)) If $\phi_{T-1} \leq m_{T-1}$: play D if the $(\tau-1)$ period history is a good history and play E otherwise.

σ^2 is defined as follows: in period 1 play C with probability π^* ,¹⁶ D with probability $1 - \pi^*$. In period two, play D if $\phi_1 > m_1$ or if the one period history is a good history and play E otherwise, where m_1 equals the value of m_{T-1} defined in equation (17) when $\pi_{T-1} = \pi^*$.

We also define:

$$\pi' = 1 - \epsilon(1 - \rho) \quad (13)$$

$$m' = \frac{\Delta C}{(\tilde{V}_H - V_L)\pi'(1 - \epsilon)} \quad (14)$$

$$\pi_{T-1} = \min \left\{ \frac{\pi^*}{1 - \epsilon(1 - \rho)}, 1 \right\} \quad (15)$$

$$m_{T-2}\pi_{T-1} = m' \quad (16)$$

$$m_{T-1} = \frac{\Delta C}{\pi_{T-1}(1 - \epsilon)\{[1 - \epsilon(1 - \rho)](V_L - C_L) - F\}} \quad (17)$$

Each strategy σ^τ has been constructed so that at any date $t \leq T-1$, the player is indifferent between playing C and D at any good history and

¹⁶Recall that π^* has been defined earlier in (8), and is the maximum probability with which α can be played in the two period game such that a player is willing to play E in the last period after receiving signal d .

also at the null history. Let us first verify this for $t \leq T - 2$. If the relevant history is a null history (which arises either if $t = 1$ or if ϕ_{t-1} is less than its critical value), a player i 's opponent j plays C with probability π' . If i 's history is a good history, j plays C for sure if j 's private history is also a good history. The probability that j is at a good history given that i is a good history is $1 - \epsilon(1 - \rho)$, which equals π' from (13). Hence in either case, the player believes that his opponent will be playing C today with probability π' . The one period gain in today's payoff from playing D as opposed to C is ΔC , while the loss in future payoffs equals $m'\pi'(1 - \epsilon)(\tilde{V}_H - V_L)$ if $t > T - 2$, and equals $m_{T-2}\pi_{T-1}\pi'(1 - \epsilon)(\tilde{V}_H - V_L)$ if $t = T - 2$.¹⁷ The definitions of m' , m_{T-2} and π_{T-1} above ensure that the today's gain equals the future loss, thus ensuring that playing C is optimal at this information set.

On the other hand, at any private history which is bad, a player believes that his opponent is playing C today with a probability which is strictly less than π' . For instance if a player has always played C , has received signal d at $t - 1$, and received signal c in all previous periods, this probability equals

$$\mu(Cc, \dots Cc, Cd) = \frac{\pi'(1 - \rho)\epsilon(1 - \epsilon)}{\pi'\epsilon + (1 - \pi')} < \pi' \quad (18)$$

The gain this period from playing D is still ΔC , but the future loss is now reduced since π' must be replaced in the expressions in the previous paragraph by $\mu(Cc, \dots Cc, Cd)$. Hence it is strictly optimal to play D at this private history. It is also easy to verify that at any other bad history, a player's belief that his opponent is playing C is less than $\mu(Cc, \dots Cc, Cd)$. For example, if a player has ever played D (or E), he knows that his opponent will play D with probability one, and hence it is strictly optimal to play D .

At date $T - 1$, we have two possibilities if $\phi_{T-2} \leq m_{T-2}$. Note that $\epsilon(1 - \rho)$ is the probability that a player's opponent will be at information set (Cc, \dots, Cd) given that a player himself is at a good history. If this is sufficiently large and greater than $1 - \pi^*$, there is no need for either player to randomize at a good history, and hence the strategy requires that C be played with probability one. However, if $1 - \epsilon(1 - \rho) < \pi^*$, players must also randomize at a good history, in order to ensure that at such a good history player i now believes that the other player plays D with probability π^* . This ensures that i has the incentive to play E in the final period in the event that he observes signal d . The definition of m_{T-1} (17) ensures that a player is indifferent

¹⁷By construction, a player is indifferent between C and D at any good history. Hence a simple way of verifying these expressions is to compute the future loss from playing D today vis-a-vis a strategy which plays C today, and D tomorrow in the event that the sunspot is less than the critical value.

between C and D at any good history at date $T - 1$, but strictly prefers D at any bad history, regardless of whether π_{T-1} equals one or is strictly smaller..

Under the equilibrium σ^T , it is easy to verify that at each date except the last one, the probability that the players play the action profile (C, C) tends to one as $\epsilon \rightarrow 0$, while the probability of playing (D, D) at date T also tends to one. Hence we have proved:

Proposition 5 *If the stage game G is repeated T times and players observe a common sunspot at the end of each period, they will be able to approximate the efficient perfect monitoring equilibrium payoff $V^*(T)$ if the noise ϵ is small.*

Since $V^*(T) \rightarrow \tilde{V}_H - C_H$ as $T \rightarrow \infty$, this proposition also implies that $\tilde{V}_H - C_H$, the symmetric efficient payoff of the stage game, can be approximated as an equilibrium payoff under imperfect private monitoring as $\epsilon \rightarrow 0$ and $T \rightarrow \infty$.

A related question is whether we can obtain the efficient average payoff $\tilde{V}_H - C_H$ as $\epsilon \rightarrow 0$ and $T \rightarrow \infty$ without recourse to a public randomization device. We conjecture that this is possible, provided that there is some discounting. As in Ellison [?] and Sekiguchi [25], one can mitigate punishments by delaying them, and by using such a mechanism it seems likely that one can achieve this (weaker) notion of efficiency.

Note that in general the strategy σ^T requires mixing in the first period, and also in period $T - 1$. This is the case even if sunspot realizations are always below their critical value, so that the game is never re-started. Randomization is required in the first period, since otherwise the other player may not have any incentive to condition behavior upon the private signals in subsequent periods. Randomization may also be required in period $T - 1$, since otherwise a player may not be willing to play distinct actions in the final period.

For computational and expositional simplicity, we have constructed an equilibrium where a player is indifferent between C and D at information sets where he is required to play C with probability one, but is strictly prefers to play D at information sets where he is required to play D . This strategy satisfies the refinement criterion set out in Bhaskar ([5]) since it is measurable with respect to a player's beliefs — a player follows the same continuation strategy when his beliefs about his opponent's continuation strategy are the same. Furthermore, at the cost of some additional complexity, one can also construct equilibria where a player has strict incentives to play C at any information set where he chooses D with probability one.

4 Concluding Comments

We now offer a brief summary of some of the related literature — some additional literature has already been discussed in previous sections. The difficulty in supporting efficient outcomes in repeated games with private signals was first pointed out by Matsushima [21], who considered pure strategy equilibria in infinitely repeated games with independent private signals. With pure strategies, the signals are uninformative, and each player's beliefs about his opponents' future actions do not change with the realization of the signal. Matsushima assumed that players adopt strategies where they do not vary their actions in response to signals unless they have a strict incentive to do so, and proved an anti-folk theorem — all pure strategy Nash equilibria satisfying this property require players to play a Nash equilibrium of the stage game in each period.

An example of equilibrium with some cooperation, in the context of a twice-repeated game, was provided by Kandori [16], and subsequently, in earlier versions of this paper. As we have discussed in section 3, the two constructions embody quite different ideas. The first example of purely non-cooperative equilibrium¹⁸ in the context of an infinitely repeated game with private signals is due to Sekiguchi [25]. Sekiguchi's work and the current paper are complementary, and both employ the idea of using mixed strategies to allow players to learn from their private signals. The idea that punishments should not be too severe also appears in Sekiguchi's work, and plays a key role in ensuring efficiency. In contrast, Compte [10] shows that if one restricts attention to grim trigger strategies in the repeated prisoners' dilemma, then payoffs must tend to the minmax level as discounting vanishes.

Our results are related to and may have implications for the work on games with imperfectly observable commitment, introduced by Bagwell [2]. Bagwell observed that the slightest amount of imperfect observation destroyed a Stackelberg leader's advantage from pre-commitment, since the Stackelberg equilibrium was no longer a pure strategy equilibrium. van Damme and Hurkens [11] showed that with one leader and one follower, the Stackelberg equilibrium could always be approximated by an equilibrium in mixed strategies, thereby ensuring that the sequential equilibrium correspondence was lower-hemicontinuous. Guth, Kirchsteiger and Ritzberger [14] show, via an example, that lower-hemicontinuity is not ensured with more than one follower.

We conclude with an analogy which may be instructive. Dynamic games

¹⁸I.e. of an equilibrium without communication, as in Compte [9] and Kandori-Matsushima [18].

where players have *private information* about past events are yet to be fully understood. At first sight, these games bear a strong resemblance to Rubinstein's [24] electronic mail game, and related infection arguments. Rubinstein's example shows that if the messages are noisy, exogenous and privately observed, players will not be able to condition their behavior on these messages.¹⁹ The repeated games we discuss differ in one respect — signals are no longer exogenous, since players may influence them via their actions. Players may adopt two devices — *individual randomization*, so that each player is uncertain about his opponent's pure strategy, and *public randomization*, which provides a favorable environment for individual randomization. These devices suffice to ensure very different results from Rubinstein's. While we are as yet far from a general theory of these games, we hope that the ideas suggested here may have a role in the development of a such a theory.

References

- [1] D. Abreu, D. Pearce and E. Stacchetti, Towards a Theory of Discounted Repeated Games with Imperfect Monitoring, *Econometrica*, 58 (1990), 1041-1064.
- [2] K. Bagwell, 1995, Commitment and Observability in Games, *Games Econ. Behav.*, 8, 271-280.
- [3] J.P. Benoit, and V. Krishna, Finitely Repeated Games, *Econometrica*, 53 (1985), 905-922.
- [4] V. Bhaskar, Repeated Games with Almost Perfect Monitoring by Privately Observed Signals, mimeo, 1994.
- [5] V. Bhaskar, The Robustness of Repeated Game Equilibria to Incomplete Payoff Information, mimeo, 2000.
- [6] V. Bhaskar, Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems, *Rev. Econ. Stud.*, 65 (1998) , 135-149.
- [7] V. Bhaskar, Notes on Communication in Finitely Repeated Games with Private Monitoring, mimeo, 1998.

¹⁹Infection arguments apply in other static contexts with imperfect information. For example, Carlsson and van Damme [8] show that a small amount of payoff uncertainty selects a unique equilibrium in games with a multiple equilibria. The analogy between repeated games with private signals and infection arguments has been explored further by Mailath and Morris [20].

- [8] H. Carlsson and E. van Damme, Global Games and Equilibrium Selection, *Econometrica*, 61 (1993), 989-1018.
- [9] O. Compte, Communication in repeated games with imperfect private monitoring, *Econometrica*, 66 (1998), 597-626.
- [10] O. Compte, On Failing to Cooperate when Monitoring is Private, mimeo, 1996.
- [11] E. van Damme, and S. Hurkens, Games with Imperfectly Observable Commitment, *Games Econ. Behav.*, 21 (1997), 282-308.
- [12] D. Fudenberg, D. Levine and E. Maskin, The Folk Theorem with Imperfect Public Information, *Econometrica*, 62 (1994), 997-1040.
- [13] E. Green, and R. Porter, Non-Cooperative Collusion under Imperfect Price Information, *Econometrica* 52 (1984), 87-100.
- [14] W. Güth, G. Kirchsteiger and K. Ritzberger, Imperfectly Observable Commitments in n-Player Games, *Games Econ. Behav.*, 23 (1998), 54-74.
- [15] J. Harsanyi, Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points, *Int. J. Game Theory* 2 (1973), 1-23.
- [16] M. Kandori, Cooperation in Finitely Repeated Games with Imperfect Private Information, mimeo, 1991.
- [17] M. Kandori, The Use of Information in Repeated Games with Imperfect Monitoring, *Rev. of Econ. Stud.*, 59 (1992), 581-593.
- [18] M. Kandori, and H. Matsushima, Private Observation, Communication and Collusion, *Econometrica*, 66 (1998), 627-652.
- [19] B. Klein and K. B. Leffler, The Role of Market Forces in Assuring Contractual Performance, *J. Polit. Economy*, 89 (1981), 615-641.
- [20] G. Mailath, and S. Morris, Repeated Games with Almost-Public Monitoring, mimeo, 1998.
- [21] H. Matsushima, On the Theory of Repeated Games with Private Information, Part I, *Econ. Letters* 35 (1991), 253-256.
- [22] R. Myerson, Multi-Stage Games with Communication, *Econometrica*, 54 (1986), 323-358.

- [23] R. Radner, R. Myerson and E. Maskin, An Example of a Repeated Partnership Game with Discounting and Uniformly Inefficient Equilibria, *Rev. Econ. Stud.*, 53 (1986), 59-70.
- [24] A. Rubinstein, The Electronic Mail Game: Strategic Behavior under “Almost Common Knowledge”, *Amer. Econ. Rev.*, 79 (1989), 385-391.
- [25] T. Sekiguchi, Efficiency in the Prisoners’ Dilemma with Private Monitoring, *J. Econ. Theory*, 76 (1997), 345-361.